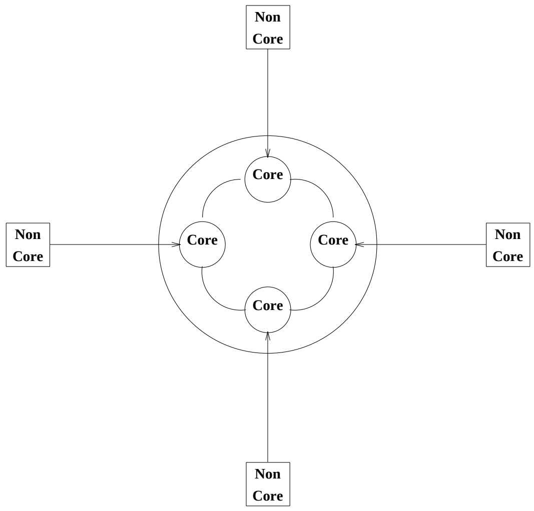
IP Routing

- Early Implementation of the internet evolved from a single backbone network, ARPANET
- ARPANET provided attached gateways to local Internet networks
  - These gateways provided the glue that held the Internet together

Routing in TCP/IP

- In the early topology a strong relationship existed between routing and network architecture
- The early structure of the Internet can be thought of as two subsets
  - A core gateway
    - Controlled by INOC
  - A non-core gateway
    - Controlled by individuals

Routing in TCP/IP



Routing in TCP/IP

- Non core gateways keep information about local destination
  - Non core gateways use default routes that point to a core gateway to send datagrams destined to other sites

Routing in TCP/IP

- Core gateways have complete knowledge of the Internet
  - No default routing among gateway routers is required
- Core gateways represent a repository of information about all possible destinations
  - Hosts can route with partial information
- Core gateways require consistent information
  - Originally produced by GGP

Gateway to Gateway Protocol GGP

- GGP is a distance vector protocol used by core gateways to exchange routing information
  - Only core gateways can participate in GGP
- GGP was designed to travel in UDP datagrams
- New core gateways can join the core
  - New core gateway get assigned core gateways neighbors
  - New core gateway informs the new neighbors about networks it can reach

Limitations of GGP

- With the rapid growth of the Internet, the topology became very complex
- Core architecture did not scale to a large size network
  - Not every site could have a core gateway connected to the backbone
  - Maintaining a consistent routing information within the core system became non-trivial
- The concept of a gateway holding complete information on an internet became unwieldy
  - As a result, GGP is no longer part of TCP/IP

Internet New Architecture Backbone Based Architecture

- Multiple backbone networks were added
  - Need for additional routing structure and protocols
  - Need for additional management tools to prevent routing loops from developing and accommodate administrative boundaries
    - This can be achieved based on the concept of ''Autonomous Domains''
    - Autonomous domains concept is crucial for discovery, propagation, validity and consistency checking of Internet routes

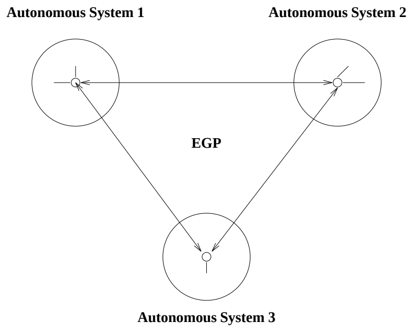
Autonomous Domains

- An autonomous system is a group of networks administered as a whole system by a single authority
  - College Campuses, Hospitals, Military installations are examples of autonomous domains
- The authority can guarantee that the routes remain consistent and viable

Autonomous Domains

- Obtain autonomous system number
- Designate one gateway to advertize reachability information to all other autonomous domains
- Advertize reachability information to the core gateway
  - Reachability information is exchanged through a specific routing protocol

Autonomous Domains



- Gateways within an autonomous domain may choose the most convenient mechanisms for discovering, propagating and checking consistency of routes

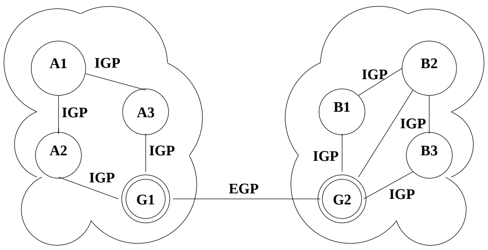
Exterior and Interior Gateways

- Based on the concept of autonomous domains, gateways were given responsibilities for only a part of an internet
  - Interior gateways exchange routing information within the same autonomous domain
    - Interior Neighbors
  - Exterior gateways exchange routing information between autonomous domains
    - Exterior Neighbors

Exterior and Interior Gateways

- Exterior Gateways use Exterior Gateway Protocols to exchange information about reachability between autonomous domains
  - Exterior Gateway Protocol (EGP) is the Internet recommended standard
- Interior Gateways use Interior Gateway Protocols to exchange routing information within an autonomous system
  - IGP does not identify a specific standard, but a concept
  - Internet does not have a leader IGP
    - Routing Information Protocol (RIP), Hello, Gated, OSPF, ..., EGP

IGP and EGP



Interior Gateways

- Interior gateways collect information about their network reachability within its autonomous system
  - Interior Gateway Protocol is needed to collect the information
- Since no early standard was available, a handful of IGPs were available
  - The most used one is Routing Information Protocol (RIP)

RIP

- RIP was developed at Berkeley based on XNS RIP
- Reached a lot of popularity before it became a standard
  - It is a straight forward implementation of distance vector routing for local networks
- RIP partitions the participants into Sets
  - Active stations, gateways, advertize their routes to others every 30 sec
  - Passive stations, hosts, listen and update their tables
    - <IP address, Distance to the network (hops)>

RIP

- A gateway keeps existing information routes until a new route of strictly lower cost is discovered
- A route becomes stale if it is not advertized again for more than 180 sec

RIP Message Formats

- Two types of messages
  - Routing information messages
  - Request information messages
- The operation field of the message specifies the command
  - 1 : Request for partial or full routing information
  - 2 : Response containing <IP Network Address, Distance> from sender's routing table
  - 3 : Turn the trace mode on
  - 4 : Turn the trace mode off
  - 5 : Reserved for Internal use

RIP Message Format

0	8	16 24 31
Command	Version	Set to 0 all the time
Family of Net 1		Set to 0 all the time
IP address of Net 1		
Set to 0 all the time		
Set to 0 all the time		
Distance to Net 1		
Family of Net 2		Set to 0 all the time
IP address to Net 2		
Set to 0 all the time		
Set to 0 all the time		
Distance to Net 2		

- Family of Net i identifies the protocol family under which the network address should be interpreted
  - RIP was general enough to support multiple protocol families
- In TCP/IP RIP travels over UDP
  - Relies on UDP to report the length of the packet (No explicit length field)

Interior Gateway Protocols

- As the Internet grew larger, it became apparent that RIP is increasingly unable to serve a large number of heterogeneous internetworks
- Two actions were taken to achieve efficient and manageable routing
  - Further divide autonomous domains into areas
  - Create a new interdomain routing protocol
    - Open Shortest Path First (OSPF)

Interior Gateway Protocols
OSPF

- OSPF has two primary characteristics
  - First, it is open, in that its specification is in the public domain
  - Second, it is based on Shortest Path First (Dijkstra Algorithm)

OSPF

- OSPF is a link state routing protocol
  - Very similar to IS0 10589
  - Developed only for the IP community
  - All fields assume addresses are exactly 4 bytes long

OSPF

- OSPF includes types of service routing
  - Routing uses both destination address and service fields to choose the best route
    - Delay, Response and Throughput are the supported service fields
- OSPF provides load balancing
  - If multiple routes are specified to the same destination at the same cost, the load is split

OSPF

- OSPF allows gateways to exchange routing information learned from other external sites
- OSPF allows virtual routing topologies
  - A virtual link between two gateways can be defined even if no physical route exists
  - Offers more flexibility

OSPF

- OSPF runs on top of IP
- Link State Advertisements (LSAs) are used to advertize link state information
  - Information on attached interfaces, metrics used, and other variables
  - OSPF routers accumulate LSAs and use SPF to compute the shortest path to each node

OSPF Routing Hierarchy
Autonomous Domains

- The largest entity within the hierarchy is the Autonomous Domain
  - OSPF is an intra-AS routing protocol
- ASs are divided into a number of ''areas''

OSPF Routing Hierarchy
Areas Domains

- An area is a group of contiguous networks and attached hosts
- Routers with multiple interfaces can participate in multiple areas
  - Referred to as ''area border routers''
  - Area border routers maintain separate topological databases for each area

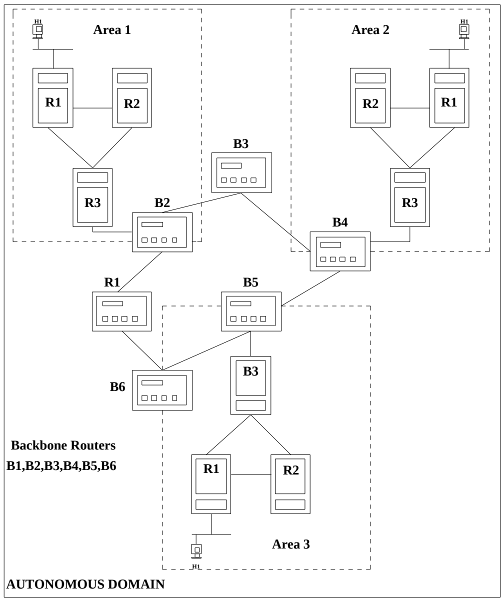
Topological Databases

- A topological database contains the collection of LSAs received from all routers in the same area
  - Essentially overall picture of networks in relationship with routers
  - Border area routers share the same LSAs, thus have identical topological database

Area Topology

- An area topology is invisible to entities outside the area
- Separate area topologies reduces the amount of routing information exchanged among routers
  - Intra-area traffic remains within the same area
  - Inter-area traffic is exchanged between areas
    - OSPF backbone is responsible for distributing routing information between areas

Hierarchical OSPF Internetworking



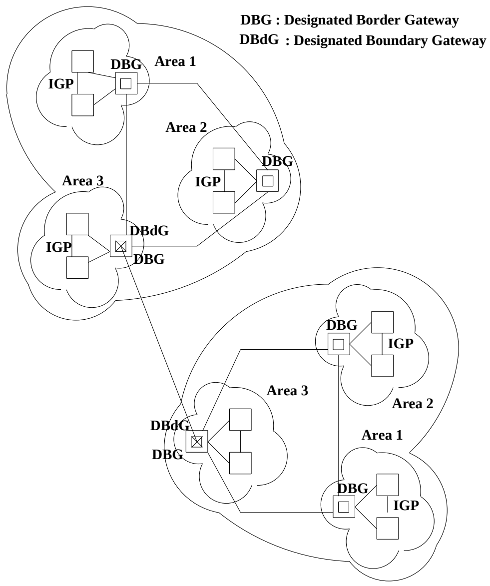
OSPF Internetworking Virtual Links

- Areas can be defined in such a way that the backbone is contiguous
  - Backbone connectivity is restored through virtual links
    - Virtual links are configured between any backbone routers that a share a link to a non-backbone router
    - Virtual links act like direct links

OSPF Algorithm

- OSPF router initializes its routing protocol data structure and waits for indication from lower layer protocols that interfaces are functional
- OSPF router uses Hello protocol to acquire neighbors, routers with interface to a common network
  - Hello packets act also as "keepalive packets"
  - On a multi-access network, supporting more than two routers, Hello message also selects designated routers
    - Designated routers are responsible for generating LSAs for the entire multi-access network

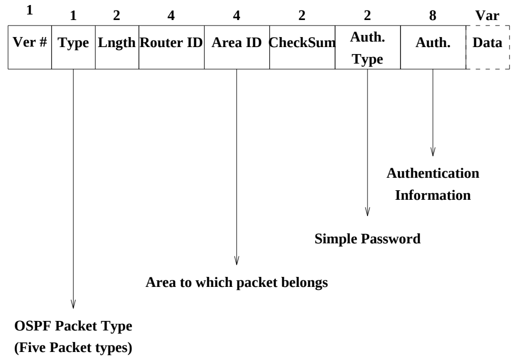
OSPF Designated Routers ASs and Areas



OSPF LSAs

- Each router periodically (or when router's state changes) sends an LSA
- LSAs serve to update Topological DataBase (TDB)
  - Based on TDB, each router computes a shortest-path tree, with itself as a root
- AS border routers running OSPF learn about exterior routes through exterior gateway protocols
  - EGP, BGP

OSPF Packet Format



OSPF Packet Type

- Hello, sent periodically to establish and maintain neighbor relationships
- Database description, used to exchange and synchronize TDB among adjacent routers
- Link state request, to request pieces of missing TDB information
  - Issued when router discovers that parts of Database description packets are not contained in its TDB
- Link state update, in response to link state requests
  - Also used for regular dissemination of four types of LSAs
- Link State Acknowledgment, to acknowledge link-state update packets

OSPF Update Packet
Four types

- Routers Links Advertisements (LSAs), describes the collected states of router's links to a specific area, for each area it belongs to
  - Flooded throughout the entire area (no further)
- Network Links Advertisements (NLAs), sent by the designated routers to describe all routers that are attached to a multi-access network
  - Flooded throughout the area containing the multi-access network

OSPF Update Packet

- Summary Links Advertisements (SLAs), to summarize routes to destination outside an area, but within the AS
  - Generated by border routers and flooded throughout the area
  - Only intra-area routers are advertized into the backbone
  - Both intra- and inter- routes are advertised into other areas
- AS External Links Advertisements (ELAs), to describe routes that is external to the AS
  - Originated by AS boundary routers
  - Only type of advertisement that is forward everywhere in the AS

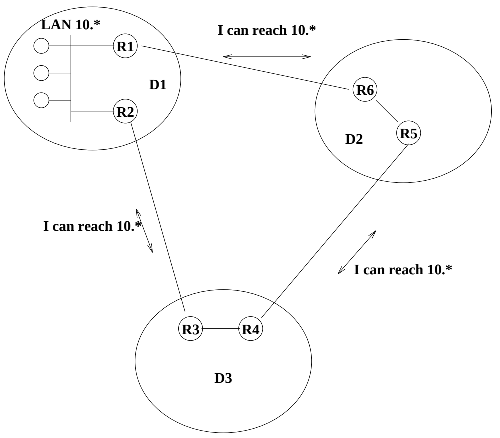
Exterior Gateway Protocol

- Exterior gateways collect information about reachability between autonomous domains
- Exterior Gateway Protocol, EGP, is a distance vector protocol used by exterior gateways to collect inter-domain routing information
  - As a first exterior gateway protocol, EGP gained widespread acceptance in the Internet

Basic Mechanisms of EGP

- Neighbor acquisition (No notion of geographic proximity)
  - Gateways request other gateways to agree to interchange reachability information
- Periodic exchange of ''still alive'' messages neighbor
- Periodic Exchange of reachability information

EGP



EGP Messages

- **Acquisition Request**
  - **A request issued by gateways to recruit neighbors**
- **Acquisition Confirm**
  - **Positive response to a recruiting request**
- **Acquisition Refuse**
  - **Negative response to a recruiting request**

EGP Messages

- **Cease Request**
  - **Request to terminates relationship**
- **Cease Confirm**
  - **Confirms a Cease request**
- **Hello, I_Heard_You,**
  - **request and response to alive messages**

EGP Messages

- **Poll request**
  - **Requests network routing update**
  Routing update
  - **Network reachability information**
- **Error**
  - **Response to incorrect message**

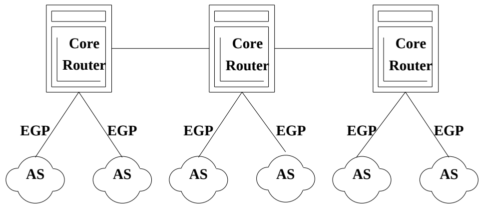
EGP Routing Update Message

Version	Type	Code (0)	Status
Checksum		Autonomous System Number	
Sequence Number		# Interior Gs	# Exterior Gs
IP Source Network			
First Gateway IP address			
# Distances			
Distance D11	# Nets at D11		
Net 1 at D11			
Net 2 at D11			
....			
Distance D12	# Nets at D12		
Net 1 At D12			
Net 2 at D12			
....			
Last Gateway IP Address			
# Distances			
...			
....			
Last Net at Last Distance for Last Gateway			

Limitations of EGP

- Third party restrictions
  - EGP restricts a non-core gateway to advertizing only networks reachable from within its autonomous domain
  - EGP does not make use of any of the distance metrics that appear in route update messages
    - Distance is only for reachability purposes. Distance are relative to source networks
    - Without routing information, only the route on which traffic flows should be advertized
    - As a result, Internet using EGP is reduced to a tree structure

EGP Internet Views



Consequences of Tree Structure

- Connectivity fails if core gateway backbone fails
- Since only one path is advertized by EGP, traffic between autonomous domains is limited to that path, even if multiple physical paths exist
- Optimal routing is difficult to achieve
- Physical failure is difficult to handle
- EGP will be eventually replaced by Border Gateway Protocol (BGP)

BGP

- BGP is a distance vector protocol that runs over TCP for reliability
- BGP does not transmit the cost to a destination
- BGP transmits the sequence of autonomous system numbers in the path to the destination IP address
- BGP routers are manually configured to enable them to rank routers according to preference.
  - A BGP update message describes the route from the BGP router that generated the update message to a destination

BGP

- Although designed as inter-AS protocol, BGP can be used within ASs
  - Two BGP routers communicating between each other must reside on the same physical network
  - Within the same AS, BGP routers build a mutually consistent view of the AS
    - BGP routers determine a router within the AS that serve as a connection point to/from certain external ASs

BGP

- Some ASs are merely passthrough for network traffic
  - Passthrough ASs carry network traffic that did not originate within and is not destined for them
    - PGB must interact with the specific intra-AS routing protocol within the passthrough AS

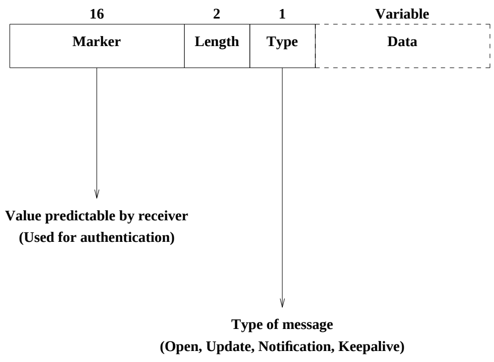
BGP

- Initial data exchange between two routers is the entire BGP routing table
  - Incremental updates are sent out as routing tables change
  - Unlike some other protocols, BGP does not require periodic refresh of the entire routing table
    - Instead, routers retain the latest version of each peer routing table

BGP

- BGP maintains a routing table with all feasible paths to a particular network
  - ○ PGB, however, advertises the primary (optimal) path in its update messages
- BGP metric, path degree of preference, are typically assigned by network administrator through configuration files
- Metric include :
  - ○ AS count (path with a smaller AS count)
  - ○ Type of link (stable, fast, reliable)

BGP Packet Format



BGP Packet Types

- Open message, sent after a transport protocol connection is established
  - ○ Open defines version number, sender AS number, hold time, authentication code, authentication data field
- Keepalive message, sent mainly to keep hold-time timers from expiring

BGP Packet Types

- Update messages, sent to provide routing updates and used to construct an AS graph
- Five attributes can characterize network paths
  - ○ Origin : IGP, EGP and incomplete (known through other means)
    - □ IGP information is preferred over EGP information (EGP fails in the presence of loops)
  - AS path, list of ASs on the path to destination
  - Next hop, IP address of the next router in path update list
  - Unreachable, route is no longer reachable
  - Inter-AS metric, provides a way for a BGP router to advertise its cost to destinations within its own AS
    - ○ Use by external routers to select optimal route into the AS

BGP Packet Types

- Notification, used to close connection between BGP routers when an error condition has been detected
  - Message header error, unacceptable message length, ...
  - Open message error, unsupported version number, unacceptable AS number or IP address, ...
  - Update message error, unacceptable update message (malformed attribute list, invalid next-hop attribute,..)
  - Hold-time expired, BGP node is declared dead

Summary

- GGP, used to provide routing information between core gateways,
- Backbone based internet, IGP and EGP
  - IGP, generic concept for protocols that provide routing information within an autonomous domain
- Intradomain domain Routing Protocol
  - RIP is a distance vector protocol
  - OSPF, created because RIP does not scale well to large heterogeneous internetworks
- Interdomain Routing Protocol
  - EGP, provides routing information between autonomous domains
  - BGP, inter-ASs routing protocol created to address most serious EGP problems